

Sum-of-squares hierarchy lower bounds for symmetric formulations*

Adam Kurpisz Samuli Leppänen Monaldo Mastrolilli
IDSIA, 6928 Manno, Switzerland {adam, samuli, monaldo}@idsia.ch

Abstract

We introduce a method for proving Sum-of-Squares (SoS)/ Lasserre hierarchy lower bounds when the initial problem formulation exhibits a high degree of symmetry. Our main technical theorem allows us to reduce the study of the positive semidefiniteness to the analysis of “well-behaved” univariate polynomial inequalities.

We illustrate the technique on two problems, one unconstrained and the other with constraints. More precisely, we give a short elementary proof of Grigoriev/Laurent lower bound for finding the integer cut polytope of the complete graph. We also show that the SoS hierarchy requires a non-constant number of rounds to improve the initial integrality gap of 2 for the MIN-KNAPSACK linear program strengthened with cover inequalities.

1 Introduction

Proving lower bounds for the Sum-of-Squares (SoS)/Lasserre hierarchy [25, 31] has attracted notable attention in the theoretical computer science community during the last decade, see e.g. [7, 12, 13, 18, 19, 27, 28, 29, 34, 36]. This is partly because the hierarchy captures many of the best known approximation algorithms based on semidefinite programming (SDP) for several natural 0/1 optimization problems (see [28] for a recent result). Indeed, it can be argued that the SoS hierarchy is the strongest candidate to be the “optimal” meta-algorithm predicted by the Unique Games Conjecture (UGC) [22, 32]. On the other hand, the hierarchy is also one of the best known candidates for refuting the conjecture since it is still conceivable that one could show that the SoS hierarchy achieves better approximation guarantees than the UGC predicts (see [6] for discussion). Despite the interest in the algorithm and due to the many technical challenges presented by semidefinite programming, only relatively few techniques are known for proving lower bounds for the hierarchy. In particular, several integrality gap results follow from applying gadget reductions to the few known original lower bound constructions.

Indeed, many of the known lower bounds for the SoS hierarchy originated in the works of Grigoriev [18, 19].¹ We defer the formal definition of the hierarchy for later and only point out that solving the hierarchy after t rounds takes $n^{O(t)}$ time. In [19] Grigoriev showed that random 3XOR or 3SAT instances cannot be solved even by $\Omega(n)$ rounds of the SoS hierarchy (some of these results were later independently rediscovered by Schoenebeck [34]). Lower bounds, such as those of [7, 36] rely on [19, 34] combined with gadget reductions. Another important lower bound was also given by Grigoriev [18] for the KNAPSACK problem (a simplified proof can be found in [20]), showing that the SoS hierarchy cannot prove within $\lfloor n/2 \rfloor$ rounds that the polytope $\{x \in [0, 1]^n : \sum_{i=1}^n x_i = n/2\}$ contains no integer point when n is odd. Using essentially the same construction as in [20], Laurent [27] independently showed that $\lfloor \frac{n}{2} \rfloor$ rounds are not

*Supported by the Swiss National Science Foundation project 200020-144491/1 “Approximation Algorithms for Machine Scheduling Through Theory and Experiments” and by Sciex Project 12.311.

¹More precisely, Grigoriev considers the positivstellensatz proof system, which is the dual of the SoS hierarchy considered in this paper. For brevity, we will use SoS hierarchy/proof system interchangeably.

enough for finding the integer cut polytope of the complete graph with n nodes, where n is odd (this result was recently shown to be tight in [15]).² By using several new ideas and techniques, but a similar starting point as in [20, 27], Meka, Potechin and Wigderson [29] were able to show a lower bound of $\Omega(\log^{1/2} n)$ for the PLANTED-CLIQUE problem. Common to the works [20, 27] and [29] is that the matrix involved in the analysis has a large kernel, and they prove that a principal submatrix is positive definite by applying the theory of association schemes [16]. It is also interesting to point out that for the class of MAX-CSPs, Lee, Raghavendra and Steurer [28] proved that the SoS relaxation yields the “optimal” approximation, meaning that SDPs of polynomial-size are equivalent in power to those arising from $O(1)$ rounds of the SoS relaxations. Then, by appealing to the result by Grigoriev/Laurent [18, 27] they showed an exponential lower bound on the size of SDP formulations for the integer cut polytope. For different techniques to obtain lower bounds, we refer for example to the recent papers [5, 23, 24] (see also Section 5.4) and the survey [13] for an overview of previous results.

In this paper we introduce a method for proving SoS hierarchy lower bounds when the initial problem formulation exhibits a high degree of symmetry. Our main technical theorem (Theorem 1) allows us to reduce the study of the positive semidefiniteness to the analysis of “well-behaved” univariate polynomial inequalities. The theorem applies whenever the solution and constraints are *symmetric*, informally meaning that all subsets of the variables of equal cardinality play the same role in the formulation (see Section 3 for the formal definition). For example, the solution in [18, 20, 27] for MAX-CUT is symmetric in this sense.

We note that exploiting symmetry reduces the number of variables involved in the analysis, and different ways of utilizing symmetry have been widely used in the past for proving integrality gaps for different hierarchies, see for example [8, 17, 19, 21, 24, 35]. An interesting difference of our approach from others is that we establish several lower bounds without fully identifying the formula of eigenvectors. More specifically, the common task in this context is to identify the spectral structure to get a simple diagonalized form. In the previous papers the moment matrices belong to the Bose-Mesner algebra of a well-studied association scheme, and hence one can use the existing theory. In this paper, instead of identifying the spectral structure completely, we identify only possible forms and propose to test all the possible candidates. This is in fact an important point, since the approach may be extended even if the underlying symmetry is imperfect or its spectral property is not well understood.

The proof of Theorem 1 is obtained by a sequence of elementary operations, as opposed to notions such as big kernel in the matrix form, the use of interlacing eigenvalues, the machinery of association schemes and various results about hyper-geometric series as in [18, 20, 27]. Thus Theorem 1 applies to the whole class of symmetric solutions, even when several conditions and machinery exploited in [18, 20, 27] cannot be directly applied. For example the kernel dimension, which was one of the important key property used to prove the results in [18, 20, 27], depends on the particular solution that is used and it is not a general property of the class of symmetric solutions. The solutions for two problems considered in this paper have completely different kernel sizes of the analyzed matrices, one large and the other zero.

We demonstrate the technique with two illustrative and complementary applications. First, we show that the analysis of the lower bound for MAX-CUT in [18, 20, 27] simplifies to few elementary calculations once the main theorem is in place. This result is partially motivated by the open question posed by O’Donnell [30] of finding a simpler proof for Grigoriev’s lower bound for the KNAPSACK problem.

As a second application we consider a constrained problem. We show that after $\Omega(\log^{1-\epsilon} n)$ levels the SoS hierarchy does not improve the integrality gap of 2 for the MIN-KNAPSACK linear program formulation strengthened with *cover inequalities* [10] introduced by Wolsey [37]. Adding cover inequalities is currently the most successful approach for capacitated covering

²The two problems, KNAPSACK and MAX-CUT in complete graphs, considered respectively in [18, 20] and in [27], are essentially the same and we will use MAX-CUT to refer to both.

problems of this type [1, 2, 3, 9, 11].

Our result is the first SoS lower bound for formulations with cover inequalities. In this application we demonstrate that our technique can also be used for suggesting the solution and for analyzing its feasibility.

Finally we point it out that the same analysis can be used to provide a non trivial lower bound to an open question raised by Laurent [26] regarding the Lasserre rank of the knapsack problem (see Section 5.4 for a discussion).

2 The SoS hierarchy

Consider a 0/1 optimization problem with $m \geq 0$ linear constraints $g_\ell(x) \geq 0$, for $\ell \in [m]$ and $x \in \mathbb{R}^n$. We are interested in approximating the convex hull of the integral points of the set $K = \{x \in \mathbb{R}^n \mid g_\ell(x) \geq 0, \forall \ell \in [m]\}$ with the SoS hierarchy defined in the following.

The form of the SoS hierarchy we use in this paper (Definition 1) is equivalent to the one used in literature (see e.g. [4, 25, 26]). It follows from applying a change of basis to the dual certificate of the refutation of the proof system [26] (see also [29] for discussion on the connection to the proof system). We use this change of basis in order to obtain a useful decomposition of the moment matrices as a sum of rank one matrices of special kind. This will play an important role in our analysis. We refer the reader to Appendix A for more details and for a mapping between the different forms.

For any $I \subseteq N = \{1, \dots, n\}$, let x_I denote the 0/1 solution obtained by setting $x_i = 1$ for $i \in I$, and $x_i = 0$ for $i \in N \setminus I$. We denote by $g_\ell(x_I)$ the value of the constraint evaluated at x_I . For each integral solution x_I , where $I \subseteq N$, in the SoS hierarchy defined below there is a variable y_I^N that can be interpreted as the “relaxed” indicator variable for the solution x_I . We point out that in this formulation of the hierarchy the number of variables $\{y_I^N : I \subseteq N\}$ is exponential in n , but this is not a problem in our context since we are interested in proving lower bounds rather than solving an optimization problem.

Let $\mathcal{P}_t(N)$ be the collection of subsets of N of size at most $t \in \mathbb{N}$. For every $I \subseteq N$, the q -zeta vector $Z_I \in \mathbb{R}^{\mathcal{P}_q(N)}$ is a 0/1 vector with J -th entry ($|J| \leq q$) equal to 1 if and only if $J \subseteq I$.³ Note that $Z_I Z_I^\top$ is a rank one matrix and the matrices considered in Definition 1 are linear combinations of these rank one matrices.

Definition 1. *The t -th round SoS hierarchy relaxation for the set K , denoted by $\text{SoS}_t(K)$, is the set of values $\{y_I^N \in \mathbb{R} : \forall I \subseteq N\}$ that satisfy*

$$\sum_{I \subseteq N} y_I^N = 1, \quad (1)$$

$$\sum_{I \subseteq N} y_I^N Z_I Z_I^\top \succeq 0, \text{ where } Z_I \in \mathbb{R}^{\mathcal{P}_{t+d}(N)} \quad (2)$$

$$\sum_{I \subseteq N} g_\ell(x_I) y_I^N Z_I Z_I^\top \succeq 0, \forall \ell \in [m], \text{ where } Z_I \in \mathbb{R}^{\mathcal{P}_t(N)} \quad (3)$$

where $d = 0$ if $m = 0$ (no linear constraints), otherwise $d = 1$.

It is straightforward to see that the SoS hierarchy formulation given in Definition 1 is a relaxation of the integral polytope. Indeed consider any feasible integral solution $x_I \in K$ and set $y_I^N = 1$ and the other variables to zero. This solution clearly satisfies Condition (1), Condition (2) because the rank one matrix $Z_I Z_I^\top$ is positive semidefinite (PSD), and Condition (3) since $x_I \in K$.

³In order to keep the notation simple, we do not emphasize the parameter q as the dimension of the vectors should be clear from the context.

3 The main technical theorem

The main result of this paper (see Theorem 1 below) allows us to reduce the study of the positive semidefiniteness for matrices (2) and (3) to the analysis of “well-behaved” univariate polynomial inequalities. It can be applied whenever the solutions and constraints are *symmetric*, namely they are invariant under all permutations π of the set N : $z_I^N = z_{\pi(I)}^N$ for all $I \subseteq N$ (equivalently when $z_I^N = z_J^N$ whenever $|I| = |J|$),⁴ where z_I^N is understood to denote either y_I^N or $g_\ell(x_I)y_I^N$. For example, the solution for MAX-CUT considered by Grigoriev [18] and Laurent [27] belongs to this class.

Theorem 1. *For any $t \in \{1, \dots, n\}$, let \mathcal{S}_t be the set of all polynomials $G_h(k) \in \mathbb{R}[k]$, for $h \in \{0, \dots, t\}$, that satisfy the following conditions:*

$$G_h(k) \in \mathbb{R}[k]_{2t} \tag{4}$$

$$G_h(k) = 0 \quad \text{for } k \in \{0, \dots, h-1\} \cup \{n-h+1, \dots, n\} \tag{5}$$

$$G_h(k) \geq 0 \quad \text{for } k \in [h-1, \dots, n-h+1] \tag{6}$$

For any fixed set of values $\{z_k^N \in \mathbb{R} : k = 0, \dots, n\}$, if the following holds

$$\sum_{k=h}^{n-h} z_k^N \binom{n}{k} G_h(k) \geq 0 \quad \forall G_h(k) \in \mathcal{S}_t \tag{7}$$

then matrix (8) is positive-semidefinite

$$\sum_{k=0}^n z_k^N \sum_{\substack{I \subseteq N \\ |I|=k}} Z_I Z_I^\top \quad (\text{where } Z_I \in \mathbb{R}^{\mathcal{P}_t(N)}) \tag{8}$$

Note that polynomial $G_h(k)$ in (6) is nonnegative in a *real interval*, and in (5) it is zero for a *finite set of integers*. Moreover, constraints (7) are trivially satisfied for $h > \lfloor n/2 \rfloor$.

Theorem 1 is actually a corollary of a technical theorem that is not strictly necessary for the applications of this paper, and therefore deferred to a later section (see Theorem 7 in Section 6). The proof (given in Section 6) is obtained by exploiting the high symmetry of the eigenvectors of the matrix appearing in (8). Condition (7) corresponds to the requirement that the Rayleigh quotient being non-negative restricted to some highly symmetric vectors (which we show are the only ones we need to consider).

4 Max-Cut for the complete graph

In the MAX-CUT problem, we are given an undirected graph and we wish to find a partition of the vertices (a cut) which maximizes the number of edges whose endpoints are on different sides of the partition (cut value). For the complete graph with n vertices, consider any solution with ω vertices on one side and the remaining $n-\omega$ on the other side of the partition. This gives a cut of value $\omega(n-\omega)$. When n is odd and for *any* $\omega \leq n/2$, Grigoriev [18] and Laurent [27] considered the following solution (reformulated in the basis considered in Definition 1, see Appendix B):

$$y_I^N = (n+1) \binom{\omega}{n+1} \frac{(-1)^{n-|I|}}{\omega - |I|} \quad \forall I \subseteq N \tag{9}$$

It is shown [18, 27] that (9) is a feasible solution for the SoS hierarchy of value $\omega(n-\omega)$, for *any* $\omega \leq n/2$ up to round $t \leq \lfloor \omega \rfloor$. In particular for $\omega = n/2$ the cut value of the SoS relaxation is

⁴We define the set-valued permutation by $\pi(I) = \{\pi(i) \mid i \in I\}$.

strictly larger than the value of the optimal integral cut (i.e. $\lfloor \frac{n}{2} \rfloor (\lfloor \frac{n}{2} \rfloor + 1)$), showing therefore an integrality gap at round $\lfloor n/2 \rfloor$.

We note that the formula for the solution (9) is essentially implied by the requirement of having exactly ω vertices on one side of the partition (see [18, 27] and [29] for more details) and the core of the analysis in [18, 27] is in showing that (9) is a feasible solution for the SoS hierarchy. By taking advantage of Theorem 1, the proof that (9) is a feasible solution for the SoS relaxation follows by observing the fact below.

Lemma 2. *For any polynomial $P(x) \in \mathbb{R}[x]$ of degree $\leq n$ and $y_i^N = y_I^N$ defined in (9) we have*

$$\sum_{k=0}^n \binom{n}{k} y_k^N P(k) = P(\omega)$$

Proof. By the polynomial remainder theorem $P(k) = (\omega - k)Q(k) + P(\omega)$, where $Q(k)$ is a unique polynomial of degree at most $n - 1$. It follows that

$$\sum_{k=0}^n \binom{n}{k} y_k^N P(k) = \underbrace{\sum_{k=0}^n \binom{n}{k} y_k^N (\omega - k)Q(k)}_{=0} + P(\omega) \underbrace{\sum_{k=0}^n \binom{n}{k} y_k^N}_{=1} = P(\omega)$$

since $\sum_{k=0}^n (-1)^k \binom{n}{k} Q(k) = 0$ for any polynomial of degree at most $n - 1$. □

Now by Lemma 2 we have $\sum_{k=0}^n y_k^N \binom{n}{k} G_h(k) = G_h(\omega)$ and the feasibility of (9) follows by Theorem 1, since we have that $G_h(\omega) \geq 0$ whenever $t \leq \omega$ for $\omega \leq n/2$.

5 Min-Knapsack with cover inequalities

The MIN-KNAPSACK problem is defined as follows: we have n items with costs c_i and profits p_i , and we want to choose a subset of items such that the sum of the costs of the selected items is minimized and the sum of the profits is at least a given demand P . Formally, this can be formulated as an integer program (IP) $\min \left\{ \sum_{j=1}^n c_j x_j : \sum_{j=1}^n p_j x_j \geq P, x \in \{0, 1\}^n \right\}$. It is easy to see that the natural linear program (LP), obtained by relaxing $x \in \{0, 1\}^n$ to $x \in [0, 1]^n$ in (IP), has an unbounded integrality gap.

By adding the *Knapsack Cover* (KC) inequalities introduced by Wolsey [37] (see also [10]), the arbitrarily large integrality gap of the natural LP can be reduced to 2 (and it is tight [10]). The KC constraints are as follows: $\sum_{j \notin A} p_j^A x_j \geq P - p(A)$ for all $A \subseteq N$, where $p(A) = \sum_{i \in A} p_i$ and $p_j^A = \min \{p_j, P - p(A)\}$. Note that these constraints are valid constraints for integral solutions. Indeed, in the “integral world” if a set A of items is picked we still need to cover $P - p(A)$; the remaining profits are “trimmed” to be at most $P - p(A)$ and this again does not remove any feasible integral solution.

The following instance [10] shows that the integrality gap implied by KC inequalities is 2: we have n items of unit costs and profits. We are asked to select a set of items in order to obtain a profit of at least $1 + 1/(n - 1)$. The resulting linear program formulation with KC inequalities is as follows (for $x_i \in [0, 1]$, $i = 1, \dots, n$)

$$(LP^+) \min \sum_{j=1}^n x_j \quad \text{s.t.} \quad \sum_{j=1}^n x_j \geq 1 + 1/(n - 1) \tag{10}$$

$$\sum_{j \in N'} x_j \geq 1 \quad \forall N' \subseteq N : |N'| = n - 1 \tag{11}$$

Note that the solution $x_i = 1/(n-1)$ is a valid fractional solution of value $1 + 1/(n-1)$ whereas the optimal integral solution has value 2. In the following we show that $\text{SoS}_t(LP^+)$, with t arbitrarily close to a logarithmic function of n , admits the same integrality gap as the initial linear program (LP^+) relaxation.

Theorem 3. *For any $\delta > 0$ and sufficiently large n' , let $t = \lfloor \log^{1-\delta} n' \rfloor$, $n = \lfloor \frac{n'}{t} \rfloor t$ and $\epsilon = o(t^{-1})$. Then the following solution is feasible for $\text{SoS}_t(LP^+)$ with integrality gap of $2 - o(1)$*

$$y_I^N = \binom{n}{|I|}^{-1} \cdot \begin{cases} \frac{(1+\epsilon)n}{(n-1)\lfloor \log n \rfloor} & \text{for } |I| = \lfloor \log n \rfloor \\ \frac{\epsilon t}{jn} & \text{for } |I| = j\frac{n}{t} \text{ and } j \in [t] \\ 1 - \sum_{\emptyset \neq I \subseteq N} y_I^N & \text{for } I = \emptyset \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

5.1 Overview of the proof

An integrality gap proof for the SoS hierarchy can in general be thought of having two steps: first, choosing a solution to the hierarchy that attains a superoptimal value, and second showing that this solution is feasible for the hierarchy. We take advantage of Theorem 1 in both steps. Here we describe the overview of our integrality gap construction while keeping the discussion informal and technical details minimal, the proof can be found in Appendix 5.2.

Choosing the solution. We make the following simplifying assumptions about the structure of the variables y_I^N : due to symmetry in the problem we set $y_I^N = y_J^N = y_k^N$ for each I, J such that $|I| = |J| = k$, and for every $I \subseteq N$ we set $y_I^N \geq 0$ in order to satisfy (2) for free. Furthermore, in order to have an integrality gap (i.e. a small objective function value), we guess that $y_0^N \approx 1$ forcing the other variables to be small due to (1).

We then show that satisfying (3) for every constraint follows from showing that

$$\sum_{k=0}^n \binom{n}{k} y_k^N (k-2) \prod_{i=1}^t (k-r_i)^2 \geq 0 \quad (13)$$

for every choice of t real variables r_i . We get this condition by observing similarities in the structure of the constraints and applying Theorem 1, then expressing the polynomial in root form.⁵ If we set $y_1^N = 0$, the only negative term in the sum corresponds to y_0^N . Then, it is clear that we need at least $t+1$ non-zero variables y_k^N , otherwise the roots r_i can be set such that the positive terms in (13) vanish and the inequality is not satisfied. Therefore, we choose exactly $t+1$ of the y_k^N to be strictly positive (and the rest 0 excluding y_0^N), and we distribute them “far away” from each other, so that no root can be placed such that the coefficient of two positive terms become small. To take this idea further, for one “very small” k' (logarithmic in n), we set $y_{k'}^N$ positive and space out the rest evenly.

Proving that the solution is feasible. We show that (13) holds for all possible r_i with our chosen solution by analysing two cases. In the first case we assume that all of the roots r_i are larger than $\log^3 n$. Then, we show that the “small” point k' we chose is enough to satisfy the condition. In the complement case, we assume that there is at least one root r_i that is smaller than $\log^3 n$. It follows that one of the evenly spaced points is “far” from any remaining root, and can be used to show that the condition is satisfied.

⁵We show that the roots r_i can be assumed to be real numbers.

5.2 Proof of Theorem 3

We start by proving the claimed integrality gap. The defined solution has an objective value that is arbitrarily close to 1 whereas the optimal integral value is 2. Indeed, the objective value of the relaxation is (see Appendix A): $\sum_{I \subseteq N} y_I^N |I| = \frac{n}{n-1}(1 + \varepsilon) + \varepsilon t \xrightarrow{n \rightarrow \infty} 1$.

The remaining part of the Theorem 3 follows by showing that the suggested solution satisfies (1), (2) and (3). Note that (1) is immediately satisfied by the definition of variables $\{y_I^N\}$, and (2) is satisfied since $y_I^N \geq 0$ and the rank one matrix $Z_I Z_I^\top$ is positive semidefinite for every $I \subseteq N$. It remains to prove that the condition (3) is also satisfied for all the constraints (10) and (11). Note that the constraint (11) is not symmetric (one variable is missing and sets of variables of the same size do not play the same role with respect to this constraint). However, the following lemma shows how to solve this issue by reducing to the form (8) of Theorem 1.

Lemma 4. *The constraint (3) holds for both (10) and (11) if the solution of Theorem 3 satisfies*

$$\sum_{\emptyset \neq I \subseteq N}^n y_I^N (|I| - 2) Z_I Z_I^\top \succeq \frac{n}{n-1} Z_\emptyset Z_\emptyset^\top \quad (14)$$

Proof. We first show that this implies that (3) holds for the constraint $\sum_{j=1}^n x_j \geq 1 + 1/(n-1)$. Since for a large n , $y_I^N = 0$ for $|I| = 1$ and $y_\emptyset^N \leq 1$, the condition (3) takes the following form

$$\sum_{\emptyset \neq I \subseteq N} y_I^N \left(|I| - \frac{n}{n-1} \right) Z_I Z_I^\top \succeq y_\emptyset^N \frac{n}{n-1} Z_\emptyset Z_\emptyset^\top$$

which is implied if (14) is satisfied (recall that $y_I^N \geq 0$). Next, we will show that (14) also implies that (3) is satisfied for the cover constraint $\sum_{j=1}^{n-1} x_j \geq 1$ (the other cases are similar). For this constraint Condition (3) can be written as

$$\begin{aligned} \sum_{\substack{n \notin I \subseteq N \\ I \neq \emptyset}} y_I^N (|I| - 1) Z_I Z_I^\top + \sum_{n \in I \subseteq N} y_I^N (|I| - 2) Z_I Z_I^\top - y_\emptyset^N Z_\emptyset Z_\emptyset^\top \\ \succeq \sum_{\substack{I \subseteq N \\ I \neq \emptyset}} y_I^N (|I| - 2) Z_I Z_I^\top - y_\emptyset^N Z_\emptyset Z_\emptyset^\top \succeq 0 \end{aligned}$$

which is also implied if (14) is satisfied. □

Now by Theorem 1, Condition (14) holds if we have $\sum_{k=1}^n y_k^N (k-2) \binom{n}{k} G_h(k) \geq \frac{n}{n-1} G_h(0)$ for $h = 0, 1, \dots, t$ and every univariate polynomial $G_h(k)$ of degree $2t$ such that $G_h(k) \geq 0$ for $k \in [h-1, \dots, n-h+1]$ and $G_h(k) = 0$ for $k \in \{0, \dots, h-1\} \cup \{n-h+1, \dots, n\}$.

Note that the only nontrivial case is for $h = 0$, since otherwise the above condition is immediately satisfied. Indeed, for $h > 0$, we have that $G_h(0) = 0$ and the only remaining terms in the sum are non-negative. Thus in order to complete the proof of Theorem 3 it is enough to show that the following is satisfied

$$\sum_{k=1}^n y_k^N (k-2) \binom{n}{k} P^2(k) \geq \frac{n}{n-1} P^2(0) \quad \forall P : \deg(P) \leq t \quad (15)$$

The following lemma (proved in Section 5.3) further reduces the interesting cases.

Lemma 5. *In order to prove that Solution (18) satisfies (15) it is sufficient to prove that (18) satisfies (15) for polynomials $P(x)$ with the following properties:*

- (a) all the roots r_1, \dots, r_t of $P(x) = 0$ are real numbers,

- (b) all the roots r_1, \dots, r_t of $P(x) = 0$ are in the range, $1 \leq r_j \leq n$ for all $j = 1, \dots, t$,
(c) the degree of $P(x)$ is exactly t .

Next we show that Solution (18) satisfies (15) and that there exists an $\varepsilon = o(t^{-1})$ as claimed.

The fundamental theorem of algebra states that any univariate polynomial of degree t has exactly t complex roots. According to Lemma 5 we can focus to polynomial with t real roots. We prove that the suggested solution satisfies (15) by expressing the generic univariate polynomial $P(k)$ using its roots r_1, \dots, r_t , so that (15) becomes

$$\sum_{k=1}^n \binom{n}{k} y_k^N (k-2) \prod_{i=1}^t (r_i - k)^2 \geq \left(1 + \frac{1}{n-1}\right) \prod_{i=1}^t r_i^2 \quad (16)$$

To show that (16) is satisfied we separate two cases: when all of the roots of the polynomial are greater or equal to a fixed threshold $\alpha = \log^3 n$ and when at least one root is smaller than this threshold. In order to simplify the computations we denote $\beta = \lfloor \log n \rfloor$.

1. $r_j \geq \alpha$ for all j . It is sufficient to show that the left-hand side term in (16) corresponding to $k = \beta$ satisfies

$$\binom{n}{\beta} y_{\beta}^N (\beta-2) \prod_{i=1}^t (r_i - \beta)^2 \geq \left(1 + \frac{1}{n-1}\right) \prod_{i=1}^t r_i^2$$

Replacing the variables with the values we get

$$\begin{aligned} \frac{n}{n-1} \frac{1+\varepsilon}{\beta} (\beta-2) \prod_{i=1}^t (r_i - \beta)^2 &\geq \frac{n}{n-1} \prod_{i=1}^t r_i^2 \\ \iff 1 + \varepsilon &\geq \prod_{i=1}^t \left(\frac{r_i}{r_i - \beta} \right)^2 \frac{1}{1 - 2\beta^{-1}} \end{aligned}$$

Since by Theorem 1 and assumption, all roots r_j satisfy $\alpha \leq r_j \leq n$. Since $\frac{r_i}{r_i - \beta} \leq \frac{\alpha}{\alpha - \beta}$ it is sufficient that it holds $1 + \varepsilon \geq \frac{1}{1 - 2\beta^{-1}} \left(\frac{\alpha}{\alpha - \beta} \right)^{2t}$.

2. There is at least one root r_j such that $r_j < \alpha$. It can be shown by straightforward induction on the number of roots that if for at least one j , $r_j < \alpha$, then there exists a point $u = l \frac{n}{t}$, $l = 1, \dots, t$ such that $\binom{n}{u} y_u > 0$ and $|u - r_i| \geq \frac{n}{2t}$ for all $i = 1, \dots, t$. Let u be such a point. It is sufficient to show that we can satisfy $\binom{n}{u} y_u^N (u-2) \prod_{i=1}^t (r_i - u)^2 \geq \frac{n}{n-1} \prod_{i=1}^t r_i^2$.

We have $\binom{n}{u} y_u^N = \frac{\varepsilon}{u}$ and the estimates $u-2 \geq \frac{u}{2}$, $(r_i - u)^2 \geq \frac{n^2}{(2t)^2}$, $\prod_{i=1}^t r_i \leq n^{t-1} \alpha$. Substituting these we get the condition $\frac{\varepsilon}{2} \left(\frac{n}{2t} \right)^{2t} \geq \frac{n}{n-1} n^{2t-2} \alpha^2$ which gives us the requirement that $\varepsilon \geq \frac{2\alpha^2}{n^2} (2t)^{2t} \frac{n}{n-1}$.

These two cases suggest that we fix ε as

$$\varepsilon = \max \left\{ \frac{1}{1 - 2\beta^{-1}} \left(1 - \frac{\beta}{\alpha} \right)^{-2t} - 1, \frac{n}{n-1} \frac{2\alpha^2}{n^2} (2t)^{2t} \right\}$$

The proof has now been reduced to showing that with this choice of ε we have $\varepsilon t \rightarrow 0$, i.e., $\varepsilon = o(t^{-1})$. Assume $\varepsilon = \frac{1}{1 - 2\beta^{-1}} \left(1 - \frac{\beta}{\alpha} \right)^{-2t} - 1$. Then $\varepsilon t = t \left(\frac{1}{1 - 2\beta^{-1}} \left(1 - \frac{\beta}{\alpha} \right)^{-2t} - 1 \right) \leq t \left(\frac{1}{1 - 2\beta^{-1}} e^{4t \frac{\beta}{\alpha}} - 1 \right)$, when $\beta/\alpha \leq 1/2$, using the estimate $1 - x \geq e^{-2x} \Rightarrow (1 - x)^{-2t} \leq e^{4xt}$ which holds when $x \leq 1/2$. Furthermore, the same estimate yields $e^x - 1 \leq 2x$ when $x \leq 1/2$. Hence, we have the bound

$$\varepsilon t \leq t \frac{1}{1 - 2\beta^{-1}} \cdot 8t \frac{\beta}{\alpha} + t \left(\frac{1}{1 - 2\beta^{-1}} - 1 \right) = \frac{8}{1 - 2\beta^{-1}} \cdot t^2 \frac{\beta}{\alpha} + \frac{2t\beta^{-1}}{1 - 2\beta^{-1}}$$

The right-hand side goes to 0 if $\frac{t^2\beta}{\alpha} \rightarrow 0$ and $\frac{t}{\beta} \rightarrow 0$ as $n \rightarrow \infty$. This is clearly the case for $t \leq \log^{1-\delta} n$, for any $\delta > 0$.

Next, assume $\varepsilon = \frac{n}{n-1} \frac{2\alpha^2}{n^2} (2t)^{2t}$. Then $\varepsilon t = t \frac{n}{n-1} \frac{2\alpha^2}{n^2} (2t)^{2t}$, which immediately yields the condition on α and t that we need $\frac{t\alpha^2}{n^2} (2t)^{2t} \rightarrow 0$ as $n \rightarrow \infty$. Substituting $\alpha = \log^3 n$ and $t = \log^{1-\delta} n$, for any $\delta > 0$, allows us to write this as $\frac{t\alpha^2}{n^2} (2t)^{2t} = \log^{1-\delta} n \frac{\log^6 n}{n^2} (2 \log^{1-\delta} n)^{2 \log^{1-\delta} n}$. By a change of variables of the form $w = \log^{1-\delta} n$ we get

$$\frac{w^{2w+\frac{7-\delta}{1-\delta}} 2^{2w}}{e^{2w\frac{1}{1-\delta}}} \leq \frac{w^{4w+\frac{7-\delta}{1-\delta}}}{e^{2w\frac{1}{1-\delta}}} = \frac{e^{(4w+\frac{7-\delta}{1-\delta}) \log w}}{e^{2w\frac{1}{1-\delta}}} = e^{(4w+\frac{7-\delta}{1-\delta}) \log w - 2w\frac{1}{1-\delta}}$$

which tends to 0 as $n \rightarrow \infty$.

5.3 Proof of Lemma 5

Lemma 5. *In order to prove that Solution (18) satisfies (15) it is sufficient to prove that (18) satisfies (15) for polynomials $P(x)$ with the following properties:*

- (a) *all the roots r_1, \dots, r_t of $P(x) = 0$ are real numbers,*
- (b) *all the roots r_1, \dots, r_t of $P(x) = 0$ are in the range, $1 \leq r_j \leq n$ for all $j = 1, \dots, t$,*
- (c) *the degree of $P(x)$ is exactly t .*

Proof. First notice that (15) is equivalent to

$$\sum_{k=1}^n y_k^N \binom{n}{k} (k-2) \prod_{j=1}^t \left(\frac{r_j - k}{r_j} \right)^2 \geq \left(1 + \frac{1}{n-1} \right) \quad (17)$$

where for the fixed n the right-hand side is constant.

- (a) Let $P(k)$ be the univariate polynomial with $2q$ complex roots (complex roots appear in conjugate pairs) i.e. $r_{2j-1} = a_j + b_j i$, $r_{2j} = a_j - b_j i$ for $j = 1, \dots, q$ and the rest real roots. Let $P'(k)$ be the polynomial with all real roots such that $r'_{2j-1} = r'_{2j} = \sqrt{a_{2j}^2 + b_{2j}^2}$ for $j = 1, \dots, q$ and $r'_j = r_j, j > 2q$.

For any $k \in N$ and $j \in [t]$, a simple calculation shows that

$$\left(\frac{r_{2j-1} - k}{r_{2j-1}} \right)^2 \left(\frac{r_{2j} - k}{r_{2j}} \right)^2 \geq \left(\frac{r'_{2j-1} - k}{r'_{2j-1}} \right)^2 \left(\frac{r'_{2j} - k}{r'_{2j}} \right)^2$$

Hence,

$$\sum_{k=1}^n y_k^N \binom{n}{k} (k-2) \prod_{j=1}^t \left(\frac{r_j - k}{r_j} \right)^2 \geq \sum_{k=1}^n y_k^N \binom{n}{k} (k-2) \prod_{j=1}^t \left(\frac{r'_j - k}{r'_j} \right)^2$$

- (b) Let $P(k)$ be the univariate polynomial with all positive roots but one i.e. $r_1 = -a$, for $a > 0$. Let $P'(k)$ be the univariate polynomial with all positive roots such that $r'_1 = a$ and $r'_j = r_j, j > 1$. Since for any $k \in N$

$$\left(\frac{-a - k}{-a} \right)^2 \geq \left(\frac{a - k}{a} \right)^2$$

and follows that,

$$\sum_{k=1}^n y_k^N \binom{n}{k} (k-2) \prod_{j=1}^t \left(\frac{r_j - k}{r_j} \right)^2 \geq \sum_{k=1}^n y_k^N \binom{n}{k} (k-2) \prod_{j=1}^t \left(\frac{r'_j - k}{r'_j} \right)^2$$

Now, let $P(k)$ be the univariate polynomial with $r_1 \in (0, 1)$ and $r_j \geq 1$, for $j > 1$. Let $P'(k)$ be the univariate polynomial with $r_1 = 1$ and $r'_j = r_j, j > 1$.

Since for any $k \in N$

$$\left(\frac{r_1 - k}{r_1} \right)^2 \geq \left(\frac{1 - k}{1} \right)^2$$

and follows that,

$$\sum_{k=1}^n y_k^N \binom{n}{k} (k-2) \prod_{j=1}^t \left(\frac{r_j - k}{r_j} \right)^2 \geq \sum_{k=1}^n y_k^N \binom{n}{k} (k-2) \prod_{j=1}^t \left(\frac{r'_j - k}{r'_j} \right)^2$$

Next, let $P(k)$ be the univariate polynomial with $r_t = an$ for $a > 1$ and $r_j \in [1, n]$, for $j \neq t$. Let $P'(k)$ be the univariate polynomial with $r_t = n$ and $r'_j = r_j, j \neq t$.

Since for any $k \in N$

$$\left(\frac{an - k}{an} \right)^2 \geq \left(\frac{n - k}{n} \right)^2$$

and follows that,

$$\sum_{k=1}^n y_k^N \binom{n}{k} (k-2) \prod_{j=1}^t \left(\frac{r_j - k}{r_j} \right)^2 \geq \sum_{k=1}^n y_k^N \binom{n}{k} (k-2) \prod_{j=1}^t \left(\frac{r'_j - k}{r'_j} \right)^2$$

- (c) Let $P(k)$ be the univariate polynomial with degree $s < t$ with all real roots. Let $P'(k)$ be the polynomial of degree t with all real roots such that $r'_j = r_j, j \leq s$ and $r'_j = n$ for $s < j \leq t$

For any $k \in N$, we have

$$1 \geq \left(\frac{n - k}{n} \right)^2$$

Hence,

$$\left(\frac{r_1 - k}{r_1} \right)^2 \cdots \left(\frac{r_s - k}{r_s} \right)^2 \geq \left(\frac{r_1 - k}{r_1} \right)^2 \cdots \left(\frac{r_s - k}{r_s} \right)^2 \left(\frac{n - k}{n} \right)^{2(t-s)}$$

and finally

$$\sum_{k=1}^n y_k^N \binom{n}{k} (k-2) \prod_{j=1}^t \left(\frac{r_j - k}{r_j} \right)^2 \geq \sum_{k=1}^n y_k^N \binom{n}{k} (k-2) \prod_{j=1}^t \left(\frac{r'_j - k}{r'_j} \right)^2$$

□

5.4 Further Results

In a recent paper [23] the authors characterize the class of the initial 0/1 relaxations that are *maximally hard* for the SoS hierarchy. Here, maximally hard means those relaxations that still have an integrality gap even after $n - 1$ rounds of the SoS hierarchy.⁶ An illustrative natural member of this class is given by the simple LP relaxation for the MIN KNAPSACK problem, i.e.

$$(LP) \min \left\{ \sum_{j=1}^n x_j : \sum_{j=1}^n x_j \geq P, x \in [0, 1]^n \right\}$$

In [23] it is shown that at level $n - 1$ the integrality gap is k , for any $k \geq 2$ if and only if $P = \Theta(k) \cdot 2^{2n}$. A natural question is to understand if the SoS hierarchy is able to reduce the gap when P is “small”.

This problem, for $P = 1/2$, was considered by Cook and Dash [14] as an example where the Lovasz-Schrijver hierarchy rank is n . Laurent [26] showed that the Sherali-Adams hierarchy rank is also equal to n and raised the open question to find the rank for the Lasserre hierarchy. She also showed that when $n = 2$, the Lasserre relaxation has an integrality gap at level 1, but leaves open whether or not this happens at level $n - 1$ for a general n . In [23] it is ruled out the possibility that the Lasserre/SoS rank is n for $n \geq 3$.

The following lemma provides a feasible solution for $\text{SoS}_t(LP)$ with integrality gap arbitrarily close to P when $t = \Omega(\log^{1-\varepsilon} n)$ and for any $P < 1$. The proof is omitted since it is similar to the proof of Theorem 3.

Theorem 6. *For any $\delta > 0$ and sufficiently large n' , let $t = \lfloor \log^{1-\delta} n' \rfloor$, $n = \lfloor \frac{n'}{t} \rfloor t$ and $\epsilon = o(t^{-1})$. Then the following solution is feasible for $\text{SoS}_t(LP^+)$ with integrality gap arbitrarily close to P .*

$$y_I^N = \binom{n}{|I|}^{-1} \cdot \begin{cases} \frac{(1+\epsilon)}{P^{\lfloor \log n \rfloor}} & \text{for } |I| = \lfloor \log n \rfloor \\ \frac{\epsilon t}{j^n} & \text{for } |I| = j \frac{n}{t} \text{ and } j \in [t] \\ 1 - \sum_{\emptyset \neq I \subseteq N} y_I^N & \text{for } I = \emptyset \\ 0 & \text{otherwise} \end{cases} \quad (18)$$

6 Proof of Theorem 1

Theorem 1 is actually a corollary of a stronger statement (see Theorem 7 below) that provides *necessary and sufficient* conditions for the matrix (8) being positive-semidefinite.

Theorem 7 uses a special family of polynomials $G_h(k) \in \mathbb{R}[k]$ whose definition is deferred to a later section (see Definition 3 in Section 6.1). We postpone the definition because it will become natural in the flow of the proof of Theorem 7. Here we remark that the polynomials $G_h(k)$ of Definition 3 satisfy the conditions (4), (5) and (6) of Theorem 1 (as shown in Lemma 14 to follow).

Theorem 7. *Let $z_k^N \in \mathbb{R}$ for $k \in \{0, \dots, n\}$. Then for any $t \in N$ the following matrix is positive-semidefinite*

$$\sum_{k=0}^n z_k^N \sum_{\substack{I \subseteq N \\ |I|=k}} Z_I Z_I^\top \quad (\text{where } Z_I \in \mathbb{R}^{\mathcal{P}_t(N)}) \quad (19)$$

if and only if

$$\sum_{k=0}^n z_k^N \binom{n}{k} G_h(k) \geq 0 \quad \text{for } h \in \{0, \dots, t\} \quad (20)$$

for every univariate polynomial $G_h(x) \in \mathbb{R}[x]$ of degree at most $2t$ as defined in Definition 3.

⁶Recall that at level n the integrality gap vanishes.

By Lemma 14, Theorem 1 is a straightforward corollary of Theorem 7. In the following we provide a proof for the latter.

6.1 Proof of Theorem 7

We study when the matrix $M = \sum_{k=0}^n z_k \sum_{I \subseteq N, |I|=k} Z_I Z_I^\top$, where $Z_I \in \mathbb{R}^{\mathcal{P}_t(N)}$ is positive-semidefinite. Theorem 7 allows us to reduce the condition $M \succeq 0$ to inequalities of the form $\sum_{k=0}^n \binom{n}{k} z_k p(k) \geq 0$, where $p(k)$ is a univariate polynomial of degree $2t$ with some additional remarkable properties.

A basic key idea that is used to obtain such a characterization is that the eigenvectors of M are “very well” structured. This structure is used to get $p(k)$ with the claimed properties.

The structure of the eigenvectors. Let Π denote the group of all permutations of the set N , i.e. the *symmetric* group. Let P_π be the permutation matrix of size $\mathcal{P}_t(N) \times \mathcal{P}_t(N)$ corresponding to any permutation π of set N , i.e. for any vector v we have $[P_\pi v]_I = v_{\pi(I)}$ for any $I \in \mathcal{P}_t(N)$ (see Footnote 4). Note that $P_\pi^{-1} = P_\pi^\top$.

Lemma 8. *For every $\pi \in \Pi$ we have $P_\pi^\top M P_\pi = M$ or, equivalently, M and P_π commute $M P_\pi = P_\pi M$.*

Proof. Let e_I denote the vector with a 1 in the I -th coordinate and 0’s elsewhere. Observe

$$P_\pi^\top Z_I = P_\pi^\top \sum_{Q \subseteq I} e_Q = \sum_{Q \subseteq I} P_\pi^\top e_Q = \sum_{Q \subseteq I} e_{\pi^{-1}(Q)} = \sum_{\pi(H) \subseteq I} e_H = \sum_{H \subseteq \pi^{-1}(I)} e_H = Z_{\pi^{-1}(I)}$$

Then, $P_\pi^\top M P_\pi = \sum_{k=0}^n z_k \sum_{I \subseteq N} P_\pi^\top Z_I Z_I^\top P_\pi = \sum_{k=0}^n z_k \sum_{I \subseteq N} Z_{\pi^{-1}(I)} Z_{\pi^{-1}(I)}^\top = M$. □

Corollary 9. *If $w \in \mathbb{R}^{\mathcal{P}_t(N)}$ is an eigenvector of M then $v = P_\pi w$ is also an eigenvector of M for any $\pi \in \Pi$.*

Proof. By the assumption, $Mw = \lambda w$ and by Lemma 8, $Mv = M(P_\pi w) = P_\pi Mw = \lambda v$. □

By using Corollary 9 we can show that the set of interesting eigenvectors have some “strong” symmetry properties that will be used in our analysis. In the simplest case, for any eigenvector w we could take the vector $u = \sum_{\pi \in \Pi} P_\pi w$ and observe that the elements of u have the form $u_I = u_J$ for each I, J such that $|I| = |J|$. If $\|u\| \neq 0$ then $u/\|u\|$ and $w/\|w\|$ are two eigenvectors corresponding to the same eigenvalue. The latter implies that by considering *only* eigenvectors having the form $u_I = u_J$ for each $|I| = |J|$ we would consider the eigenvalue corresponding to the “unstructured” eigenvector w as well. This is not the case in general, however, since it is possible that $\sum_{\pi \in \Pi} P_\pi w = 0$.

We overcome this obstacle by restricting the permutations in a way which guarantees u to be non-zero. Before going into the details, we introduce some notation.

Definition 2. *For any $H \subseteq N$, we denote by Π_H the permutation group that fixes the set H in the following sense: $\pi \in \Pi_H \Leftrightarrow \pi(H) = H$.*

Note that the definition is equivalent to saying that $\pi \in \Pi_H$ if and only if $\pi(i) \in H$ for every $i \in H$ and $\pi(i) \notin H$ for every $i \notin H$.

Now, we choose a subset $H \subseteq N$ such that $\sum_{\pi \in \Pi_H} P_\pi w = 0$ for each I such that $|I| < |H|$ and $u = \sum_{\pi \in \Pi_H} P_\pi w \neq 0$. Such a set H always exists, since otherwise w is a zero vector, since if there is one non-zero entry w_J in w , we can take $H = J$ and the resulting u is non-zero. The choice of H is not unique, but we can always assume that it is the subset of the first $h = |H|$

elements from N , i.e. $H = \{1, \dots, h\}$. Indeed, if it is not the case, there exists a permutation $\pi \in \Pi$ that maps H to the subset of the first $|H|$ elements from N and such that $P_\pi w$ is an eigenvector of M by Lemma 9. Now it holds that $u \neq 0$ and the vector $u/\|u\|$ is a unit eigenvector corresponding to the same eigenvalue as w and has many elements that are equal to each other.

Lemma 10. *Let $w \in \mathbb{R}^{P_t(N)}$ be a unit eigenvector of M corresponding to eigenvalue λ , and H be the smallest subset of N such that $u = \sum_{\pi \in \Pi_H} P_\pi w \neq 0$. Then $u/\|u\|$ is also a unit eigenvector of M corresponding to eigenvalue λ .*

The following lemma shows the structure of eigenvectors obtained from summing the permutations of any “unstructured” eigenvector.

Lemma 11. *Let $u = \sum_{\pi \in \Pi_H} P_\pi w$. Then the vector u is invariant under the permutations of Π_H , namely $u_I = u_{\pi(I)}$ for $\pi \in \Pi_H$. Equivalently, $u_I = u_J$ for all $|I| = |J|$ such that $|I \cap H| = |J \cap H|$.*

Proof. We have $u_I = \left[\sum_{\pi \in \Pi_H} P_\pi w \right]_I = \sum_{\pi \in \Pi_H} [P_\pi w]_I = \sum_{\pi \in \Pi_H} w_{\pi(I)} = \sum_{\pi \in \Pi_H} w_{\pi(\pi(I))} = u_{\pi(I)}$ where the last but one equality follows since permutations are bijective. The claim follows by observing that for all $|I| = |J|$ such that $|I \cap H| = |J \cap H|$ there exists $\pi \in \Pi_H$ such that $\pi(I) = J$. □

Lemma 10, Lemma 11 and the arguments above imply Lemma 12.

Lemma 12. *For any eigenvalue λ of M there exists an $h = 0, 1, \dots, t$ such that the following is an eigenvector corresponding to λ :*

$$u_h = \sum_{i=0}^t \sum_{j=0}^{\min\{h,i\}} \alpha_{i,j} b_{i,j} \quad (21)$$

where $H = \{1, \dots, h\}$, $\alpha_{i,j} \in \mathbb{R}$ and $b_{i,j} \in \mathbb{R}^{P_t(N)}$ such that $[b_{i,j}]_Q = 1$ if $|Q| = i$ and $|Q \cap H| = j$, $[b_{i,j}]_Q = 0$ otherwise.

By Lemma 12, we have that the positive semidefiniteness condition of M follows by ensuring that for any $h = 0, 1, \dots, t$ we have $u_h^\top M u_h \geq 0$, i.e.

$$u_h^\top M u_h = \sum_{k=0}^n z_k \underbrace{\sum_{\substack{I \subseteq N \\ |I|=k}} \left(u_h^\top Z_I \right)^2}_{A_k} = \sum_{k=0}^n z_k \sum_{\substack{I \subseteq N \\ |I|=k}} \left(\sum_{i=0}^t \sum_{j=0}^{\min\{h,i\}} \alpha_{i,j} b_{i,j}^\top Z_I \right)^2 \geq 0$$

In the following (Lemma 13) we show that the above values A_k are interpolated by the univariate polynomial $G_h(x)$ defined in Definition 3. In Lemma 14 we prove some remarkable properties of $G_h(x)$ as claimed in Theorem 1.

Definition 3. *For any $h \in \{0, \dots, t\}$, let $G_h(k) \in \mathbb{R}[k]$ be a univariate polynomial defined as follows*

$$G_h(k) = \sum_{r=0}^h \binom{h}{r} h_r(k) \left(\sum_{j=0}^h \binom{r}{j} p_j(k-r) \right)^2 \quad (22)$$

where $h_r(k) = k^r \cdot (n-k)^{h-r}$ and $p_j(k-r) = \sum_{i=0}^{t-j} \alpha_{i+j,j} \binom{k-r}{i}$ (for $\alpha_{i,j} \in \mathbb{R}$).⁷

⁷Denote by $x^{\underline{m}} = x(x-1) \cdots (x-m+1)$ the falling factorial (with the convention that $x^{\underline{0}} = 1$).

Lemma 13. For every $k = 0, \dots, n$ the following identity holds $A_k = \binom{n}{k} \frac{1}{n!} G_h(k)$.

Proof. We start noting that for every $i = 0, \dots, t$, $j = 0, \dots, |H|$ we have ⁸

$$b_{i,j}^\top Z_I = \binom{|I \cap H|}{j} \binom{|I \setminus H|}{i-j}$$

Indeed

$$b_{i,j}^\top Z_I = \sum_{Q \in \mathcal{P}_t(N)} (b_{i,j})_Q (z_I)_Q = \sum_{\substack{Q \subseteq I, |Q|=i \\ |Q \cap H|=j}} 1 = \binom{|I \cap H|}{j} \binom{|I \setminus H|}{i-j}$$

It follows that we have

$$\begin{aligned} \sum_{\substack{I \subseteq N \\ |I|=k}} (u^\top Z_I)^2 &= \sum_{\substack{I \subseteq N \\ |I|=k}} \left(\sum_{i=0}^t \sum_{j=0}^{|H|} \alpha_{i,j} b_{i,j}^\top Z_I \right)^2 \\ &= \sum_{\substack{I \subseteq N \\ |I|=k}} \left(\sum_{i=0}^t \sum_{j=0}^{|H|} \alpha_{i,j} \binom{|I \cap H|}{j} \binom{|I \setminus H|}{i-j} \right)^2 \end{aligned}$$

Splitting the sum over I by considering the intersections $I \cap H$ of sizes $r = 0, \dots, |H|$, we have

$$\begin{aligned} &\sum_{r=0}^{|H|} \sum_{\substack{|I|=k \\ |I \cap H|=r}} \left(\sum_{i=0}^t \sum_{j=0}^{|H|} \alpha_{i,j} \binom{r}{j} \binom{k-r}{i-j} \right)^2 \\ &= \sum_{r=0}^{|H|} \binom{|H|}{r} \binom{n-|H|}{k-r} \left(\sum_{j=0}^{|H|} \binom{r}{j} \sum_{i=0}^t \alpha_{i,j} \binom{k-r}{i-j} \right)^2 \end{aligned}$$

Finally, we shift the sum over i by j and thus justify the equality

$$\sum_{\substack{I \subseteq N \\ |I|=k}} (u^\top Z_I)^2 = \sum_{r=0}^{|H|} \binom{|H|}{r} \binom{n-|H|}{k-r} \left(\sum_{j=0}^{|H|} \binom{r}{j} \sum_{i=0}^{t-j} \alpha_{i+j,j} \binom{k-r}{i} \right)^2$$

Now, the sum over i is a Newton polynomial that we denote by $p_j(k-r) = \sum_{i=0}^{t-j} \alpha_{i+j,j} \binom{k-r}{i}$. Note that by definition, here $\deg(p) = t-j$. Furthermore, observe that

$$\binom{n-|H|}{k-r} = \binom{n}{k} \frac{1}{n^{|H|}} k^x \cdot (n-k)^{|H|-r}$$

and writing $h_r(k) = k^x \cdot (n-k)^{|H|-r}$ yields the claim. \square

It follows that for any unit eigenvector u of the form (21) the corresponding eigenvalue is equal to $u^\top M u = \frac{1}{n!} \sum_{k=0}^n z_k \binom{n}{k} G_h(k)$. Theorem 7 requires that $\sum_{k=0}^n z_k \binom{n}{k} G_h(k) \geq 0$ which implies that the eigenvalue $u^\top M u$ is nonnegative. In the following section we complete the proof by showing that the polynomials $G_h(k)$ of Definition 3 satisfy the conditions (4), (5) and (6) of Theorem 1 (as shown in Lemma 14).

⁸Recall that $\binom{n}{-k} = \binom{n}{n+k} = 0$ for any positive integer k .

6.2 Properties of the univariate polynomials

Lemma 14. *For any $h \in \{0, \dots, t\}$, the polynomials $G_h(k)$ as defined in Definition 3 have the following properties:*

- (a) $G_h(k)$ is a univariate polynomial of degree at most $2t$,
- (b) $G_h(k) \geq 0$ for $k \in [h-1, n-h+1]$
- (c) $G_h(k) = 0$ for every $k \in \{0, \dots, h-1\} \cup \{n-h+1, \dots, n\}$.

Proof of (a).

$$\begin{aligned}
G_h(k) &= \sum_{r=0}^h \binom{h}{r} h_r(k) \left(\sum_{j=0}^h \binom{r}{j} p_j(k-r) \right)^2 \\
&= \sum_{r=0}^h \binom{h}{r} h_r(k) \left(\sum_{i=0}^h \sum_{j=0}^h \binom{r}{i} \binom{r}{j} p_i(k-r) p_j(k-r) \right) \\
&= \sum_{r=0}^h \binom{h}{r} h_r(k) \left(\sum_{i=0}^h \sum_{j=0}^h \binom{r}{i} \binom{r}{j} \left(\sum_{a=0}^{t-i} \alpha_{a+i,i} \binom{k-r}{a} \right) \left(\sum_{b=0}^{t-j} \alpha_{b+j,j} \binom{k-r}{b} \right) \right) \\
&= \sum_{i=0}^h \sum_{j=0}^h \sum_{a=0}^{t-i} \sum_{b=0}^{t-j} \alpha_{a+i,i} \alpha_{b+j,j} \underbrace{\sum_{q=0}^a \sum_{s=0}^b \binom{k-h}{q} \binom{k-h}{s} \left(\sum_{r=0}^h \binom{h}{r} h_r(k) \binom{r}{i} \binom{r}{j} \binom{h-r}{a-q} \binom{h-r}{b-s} \right)}_{B(k)}
\end{aligned}$$

Note that $\binom{k-r}{a} = \sum_{q=0}^a \binom{k-h}{q} \binom{h-r}{a-q}$ by Vandermonde's identity. We prove the theorem by showing that $B(k)$ has degree not larger than $2t$.

$$B(k) = \binom{k-h}{q} \binom{k-h}{s} \underbrace{\left(\sum_{r=0}^h \binom{h}{r} k^x (n-k)^{h-r} \overbrace{\binom{r}{i} \binom{r}{j} \binom{h-r}{a-q} \binom{h-r}{b-s}}^{f(r)} \right)}_{C(k)}$$

By Lemma 15 below, the degree of $C(k)$ is at most $i+j+a-q+b-s$ and thus the degree of $B(k)$ is at most $i+j+a+b=2t$

Lemma 15. *The degree of $C(k)$ is at most $i+j+a-q+b-s$.*

The claim follows by showing that the degree of $C(k)$ is at most the degree of $f(r)$. The degree of $f(r)$ is $i+j+a-q+b-s$.

Recall that the forward difference of function $g(X)$ with respect to variable X is a finite difference defined by $\Delta_X[g(X)] = g(X+1) - g(X)$. Higher order differences are obtained by repeated operations of the forward difference operator. We will use $\Delta_X^\ell[g(X)]_{X=b}$ to denote the ℓ -th forward difference evaluated at $X=b$. We will use the following easy to check identity: $\Delta_X^d[(k+X)^{\overline{r+d}}] = (k+X)^{\overline{r}}(r+d)^{\underline{d}}$.

First note that any polynomial $f(r)$ of degree δ can be written as linear combinations of polynomials $P_d(r) = (r+1)^{\overline{d}} = (r+d)^{\underline{d}}$ with $0 \leq d \leq \delta$. It follows that the claim follows by showing

that the degree of the following $C'(k)$ is at most the degree of $P_d(r)$

$$\begin{aligned}
C'(k) &= \sum_{r=0}^h \binom{h}{r} (n-k)^{\underline{h-r}} \cdot k^r (r+d)^{\underline{d}} \\
&= \sum_{r=0}^h \binom{h}{r} (n-k)^{\underline{h-r}} \cdot \Delta_X^d \left[(k+X)^{\underline{r+d}} \right]_{X=0} \\
&= \Delta_X^d \left[\sum_{r=0}^h \binom{h}{r} (n-k)^{\underline{h-r}} (k+X)^{\underline{r+d}} \right]_{X=0} \\
&= \Delta_X^d \left[(k+X)^{\underline{d}} \sum_{r=0}^h \binom{h}{r} (n-k)^{\underline{h-r}} (k+X-d)^{\underline{r}} \right]_{X=0} \\
&= \Delta_X^d \left[(k+X)^{\underline{d}} (n+X-d)^{\underline{h}} \right]_{X=0}
\end{aligned}$$

where we have used the linearity of the forward difference operator and the Vandermonde's identity to derive the last equality. The claim follows by observing that the forward difference operator does not increase the degree of its argument and therefore $C'(k)$ has degree at most d .

Proof of (b). Let $k \in [h-1, n-h+1]$. We have

$$G_h(k) = \sum_{r=0}^h \binom{h}{r} h_r(k) \left(\sum_{j=0}^h \binom{r}{j} p_j(k-r) \right)^2$$

where $h_r(k) = k^r \cdot (n-k)^{\underline{h-r}} \geq 0$ for each $r = 0, \dots, h$. Therefore $G_h(k)$ is a sum of non-negative numbers $\left(\sum_{j=0}^h \binom{r}{j} p_j(k-r) \right)^2$ weighted by positive coefficients $h_r(k)$.

Proof of (c). From Lemma 13 we have that

$$\frac{1}{n^{\underline{h}}} \binom{n}{k} G_h(k) = \sum_{\substack{I \subseteq N \\ |I|=k}} \left(u^\top Z_I \right)^2$$

Therefore $G_h(k) = 0$ for $k \in \{0, \dots, h-1\} \cup \{n-h+1, \dots, n\}$ if we can show that $u^\top Z_Q = 0$ for all $Q \subseteq N$ such that $|Q| = k$.

With this aim, we start noting that for every set $S \subseteq N$ we have that the permutation group Π_S is the same as $\Pi_{N \setminus S}$. Moreover if $|S| < h$ then $\sum_{\pi \in \Pi_S} P_\pi u = 0$, otherwise we would obtain a set S smaller than H with $\sum_{\pi \in \Pi_S} P_\pi u \neq 0$ (contradicting our assumption that H is a set with the smallest set size having $\sum_{\pi \in \Pi_H} P_\pi u \neq 0$).

Now consider any set I such that $I \subseteq Q$ with $Q \in \{S, N \setminus S\}$ and $|S| < h$. By the previous observations it follows that $[\sum_{\pi \in \Pi_Q} P_\pi u]_I = \sum_{\pi \in \Pi_Q} P_\pi u_I = \sum_{\pi \in \Pi_Q} u_{\pi(I)} = 0$. Note that since $I \subseteq Q$ the set $\{\pi(I) : \pi \in \Pi_Q\}$ is equal to $\{J : J \subseteq Q, |J| = |I|\}$, since Π_Q is the permutation group that maps any element I from Q to any other element from Q of the same size. It follows that $\sum_{\pi \in \Pi_Q} u_{\pi(I)} = \sum_{J \subseteq Q, |J|=|I|} u_J = 0$. Using the latter we get

$$u^\top Z_Q = \sum_{J \subseteq Q} u_J = \sum_{i=0}^{|Q|} \sum_{J \subseteq Q, |J|=i} u_J = 0$$

proving the claim.

Acknowledgements. The authors would like to express their gratitude to Ola Svensson for helpful discussions and ideas regarding this paper.

References

- [1] N. Bansal, N. Buchbinder, and J. Naor. Randomized competitive algorithms for generalized caching. In *STOC*, pages 235–244, 2008.
- [2] N. Bansal, A. Gupta, and R. Krishnaswamy. A constant factor approximation algorithm for generalized min-sum set cover. In *SODA*, pages 1539–1545, 2010.
- [3] N. Bansal and K. Pruhs. The geometry of scheduling. In *FOCS*, pages 407–414, 2010.
- [4] B. Barak, F. G. S. L. Brandão, A. W. Harrow, J. A. Kelner, D. Steurer, and Y. Zhou. Hypercontractivity, sum-of-squares proofs, and their applications. In *STOC*, pages 307–326, 2012.
- [5] B. Barak, S. O. Chan, and P. Kothari. Sum of squares lower bounds from pairwise independence. In *STOC*, 2015.
- [6] B. Barak and D. Steurer. Sum-of-squares proofs and the quest toward optimal algorithms. *Electronic Colloquium on Computational Complexity (ECCC)*, 21:59, 2014.
- [7] A. Bhaskara, M. Charikar, A. Vijayaraghavan, V. Guruswami, and Y. Zhou. Polynomial integrality gaps for strong sdp relaxations of densest k -subgraph. In *SODA*, pages 388–405, 2012.
- [8] G. Blekherman, J. ao Gouveia, and J. Pfeiffer. Sums of squares on the hypercube. *CoRR*, abs/1402.4199, 2014.
- [9] T. Carnes and D. B. Shmoys. Primal-dual schema for capacitated covering problems. In *IPCO*, pages 288–302, 2008.
- [10] R. D. Carr, L. Fleischer, V. J. Leung, and C. A. Phillips. Strengthening integrality gaps for capacitated network design and covering problems. In *SODA*, pages 106–115, 2000.
- [11] D. Chakrabarty, E. Grant, and J. Könemann. On column-restricted and priority covering integer programs. In *IPCO*, pages 355–368, 2010.
- [12] K. K. H. Cheung. Computation of the Lasserre ranks of some polytopes. *Mathematics of Operation Research*, 32(1):88–94, 2007.
- [13] E. Chlamtac and M. Tulsiani. Convex relaxations and integrality gaps. In *Handbook on Semidefinite, Conic and Polynomial Optimization*, volume 166, pages 139–169. Springer, 2011.
- [14] W. Cook and S. Dash. On the matrix-cut rank of polyhedra. *Mathematics of Operations Research*, 26(1):19–30, 2001.
- [15] H. Fawzi, J. Saunderson, and P. Parrilo. Sparse sum-of-squares certificates on finite abelian groups. *CoRR*, abs/1503.01207, 2015.
- [16] C. Godsil. Association schemes. Lecture Notes available at <http://quoll.uwaterloo.ca/mine/Notes/assoc2.pdf>, 2010.
- [17] M. X. Goemans and L. Tunçel. When does the positive semidefiniteness constraint help in lifting procedures? *Math. Oper. Res.*, 26(4):796–815, 2001.
- [18] D. Grigoriev. Complexity of positivstellensatz proofs for the knapsack. *Computational Complexity*, 10(2):139–154, 2001.
- [19] D. Grigoriev. Linear lower bound on degrees of positivstellensatz calculus proofs for the parity. *Theoretical Computer Science*, 259(1-2):613–622, 2001.
- [20] D. Grigoriev, E. A. Hirsch, and D. V. Pasechnik. Complexity of semi-algebraic proofs. In *STACS*, pages 419–430, 2002.
- [21] S. Hong and L. Tunçel. Unification of lower-bound analyses of the lift-and-project rank of combinatorial optimization polyhedra. *Discrete Applied Mathematics*, 156(1):25–41, 2008.
- [22] S. Khot. On the power of unique 2-prover 1-round games. In *STOC*, pages 767–775, 2002.

- [23] A. Kurpisz, S. Leppänen, and M. Mastrolilli. A lasserre lower bound for the min-sum single machine scheduling problem. In *ESA*, pages 853–864, 2015.
- [24] A. Kurpisz, S. Leppänen, and M. Mastrolilli. On the hardest problem formulations for the 0/1 lasserre hierarchy. In *ICALP*, pages 872–885, 2015.
- [25] J. B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001.
- [26] M. Laurent. A comparison of the Sherali-Adams, Lovász-Schrijver, and Lasserre relaxations for 0-1 programming. *Mathematics of Operations Research*, 28(3):470–496, 2003.
- [27] M. Laurent. Lower bound for the number of iterations in semidefinite hierarchies for the cut polytope. *Math. Oper. Res.*, 28(4):871–883, 2003.
- [28] J. R. Lee, P. Raghavendra, and D. Steurer. Lower bounds on the size of semidefinite programming relaxations. In *STOC*, pages 567–576, 2015.
- [29] R. Meka, A. Potechin, and A. Wigderson. Sum-of-squares lower bounds for planted clique. In *STOC*, pages 87–96, 2015.
- [30] R. O’Donnell. Approximability and proof complexity. Talk at ELC Tokyo. Slides available at <http://www.cs.cmu.edu/~odonnell/slides/approx-proof-cxty.pps>, 2013.
- [31] P. Parrilo. *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization*. PhD thesis, California Institute of Technology, 2000.
- [32] P. Raghavendra. Optimal algorithms and inapproximability results for every csp? In *STOC*, pages 245–254, 2008.
- [33] T. Rothvoß. The lasserre hierarchy in approximation algorithms. Lecture Notes for the MAPSP 2013 - Tutorial, June 2013.
- [34] G. Schoenebeck. Linear level Lasserre lower bounds for certain k-csps. In *FOCS*, pages 593–602, 2008.
- [35] T. Stephen and L. Tunçel. On a representation of the matching polytope via semidefinite liftings. *Math. Oper. Res.*, 24(1):1–7, 1999.
- [36] M. Tulsiani. Csp gaps and reductions in the Lasserre hierarchy. In *STOC*, pages 303–312, 2009.
- [37] L. A. Wolsey. Facets for a linear inequality in 0–1 variables. *Mathematical Programming*, 8:168–175, 1975.

Appendix

A The SoS hierarchy

In this section we recall the usual definition of the SoS/Lasserre hierarchy [25] and justify Definition 1. Notice that the SDP hierarchy that we discuss here is the dual certificate of a refutation of the positivstellensatz proof system, for further information about the connection to the proof system we refer the reader to [29]. In our setting we restrict ourselves to problems with 0/1-variables and linear constraints. More precisely, we consider the following general optimization problem \mathbb{P} : given a multilinear polynomial $f : \{0, 1\}^n \rightarrow \mathbb{R}$

$$\mathbb{P} : \quad \min\{f(x) \mid x \in \{0, 1\}^n \cap K\} \tag{23}$$

where K is a polytope defined by m linear inequalities $g_\ell(x) \geq 0$ for $\ell \in [m]$. Many basic optimization problems are special cases of \mathbb{P} . For example, any k -ary boolean constraint satisfaction problem, such as MAX-CUT, is captured by (23) where a degree k function $f(x)$ counts the number of satisfied constraints, and no linear constraints $g_\ell(x) \geq 0$ are present. Also any 0/1 integer linear program is a special case of (23), where $f(x)$ is a linear function.

Lasserre [25] proposed a hierarchy of SDP relaxations parameterized by an integer r ,

$$\min\{L(f) \mid L : \mathbb{R}[X]_{2r} \rightarrow \mathbb{R}, L(1) = 1, L(x^2 - x) = 0 \text{ and } L(u^2), L(u^2 g_\ell) \geq 0, \forall \text{ polynomial } u\} \quad (24)$$

where $L : \mathbb{R}[X]_{2r} \rightarrow \mathbb{R}$ is a linear map with $\mathbb{R}[X]_{2r}$ denoting the ring $\mathbb{R}[X]$ restricted to polynomials of degree at most $2r$.⁹ Note that (24) is a relaxation since one can take L to be the evaluation map $f \rightarrow f(x^*)$ for any optimal solution x^* .

Relaxation (24) can be equivalently formulated in terms of *moment matrices* [25]. In the context of this paper, this matrix point of view is more convenient to use and it is described below. In our notation we mainly follow the survey of Laurent [26] (see also [33]).

Variables and Moment Matrix. Let N denote the set $\{1, \dots, n\}$. The collection of all subsets of N is denoted by $\mathcal{P}(N)$. For any integer $t \geq 0$, let $\mathcal{P}_t(N)$ denote the collection of subsets of N having cardinality at most t . Let $y \in \mathbb{R}^{\mathcal{P}(N)}$. For any nonnegative integer $t \leq n$, let $M_t(y)$ denote the matrix with (I, J) -entry $y_{I \cup J}$ for all $I, J \in \mathcal{P}_t(N)$. Matrix $M_t(y)$ is termed in the following as the *t-moment matrix* of y . For a linear function $g(x) = \sum_{i=1}^n g_i \cdot x_i + g_0$, we define $g * y$ as a vector, often called *shift operator*, where the I -th entry is $(g * y)_I = \sum_{i=1}^n g_i y_{I \cup \{i\}} + g_0 y_I$. Let f denote the vector of coefficients of polynomial $f(x)$ (where f_I is the coefficient of monomial $\prod_{i \in I} x_i$ in $f(x)$).

Definition 4. The t -th round SoS (or Lasserre) relaxation of problem (23), denoted as $\text{SoS}_t(\mathbb{P})$, is the following

$$\text{SoS}_t(\mathbb{P}) : \quad \min \left\{ \sum_{I \subseteq N} f_I y_I \mid y \in \mathbb{R}^{\mathcal{P}_{2t+2d}(N)} \text{ and } y \in \mathbb{M} \right\} \quad (25)$$

where \mathbb{M} is the set of vectors $y \in \mathbb{R}^{\mathcal{P}_{2t+2d}(N)}$ that satisfy the following PSD conditions

$$y_\emptyset = 1, \quad (26)$$

$$M_{t+d}(y) \succeq 0, \quad (27)$$

$$M_t(g_\ell * y) \succeq 0 \quad \ell \in [m] \quad (28)$$

where $d = 0$ if $m = 0$ (no linear constraints), otherwise $d = 1$.

Change of variables. A solution of the SoS hierarchy as defined in Definition 4 is given by a vector $y \in \mathbb{R}^{\mathcal{P}_{2t+2d}(N)}$. Next we show we can make a change of basis and replace the variables y_I with other variables y_I^N that are indexed by *all* the subsets of N . Variable y_I^N can be interpreted as the “relaxed” indicator variable for the integral solution x_I , i.e. the 0/1 solution obtained by setting $x_i = 1$ for $i \in I$, and $x_i = 0$ for $i \in N \setminus I$. We use this change of basis in order to obtain a useful decomposition of the moment matrix as a sum of rank one matrices of special kind. Here it is not necessary to distinguish between the moment matrix of the variables and constraints, hence in what follows we denote a generic vector by $w \in \mathbb{R}^{\mathcal{P}_{2q}(N)}$, where q is either t or $t + 1$.

Definition 5. Let $w \in \mathbb{R}^{\mathcal{P}_{2q}(N)}$. For every $I \in \mathcal{P}(N)$, define a vector $w^N \in \mathbb{R}^{\mathcal{P}(N)}$ such that

$$w_I = \sum_{I \subseteq H \subseteq N} w_H^N$$

⁹In [4], $L(p)$ is written $\tilde{\mathbb{E}}[p]$ and called the “pseudo-expectation” of p .

Note that the inverse (for $|I| \leq 2t$) is

$$w_I^N = \sum_{H \subseteq N \setminus I, |H \cup I| \leq 2t} (-1)^{|H|} w_{I \cup H} \quad (29)$$

To simplify the notation, we note that the moment matrix of the variables is structurally similar to the moment matrix of the constraints: if $z \in \mathbb{R}^{\mathcal{P}_{2q}(N)}$ is a vector such that $z_I = \sum_{i=1}^N A_{\ell i} y_{I \cup \{i\}} - b_{\ell} y_I$ for some ℓ , then $[M_t(g_{\ell} * y)]_{I,J} = z_{I \cup J}$. Hence, the following lemma holds for the moment matrix of variables and constraints.

Lemma 16. *Let $w \in \mathbb{R}^{\mathcal{P}_{2q}(N)}$, and $M \in \mathbb{R}^{\mathcal{P}_q(N) \times \mathcal{P}_q(N)}$ such that $M_{I,J} = w_{I \cup J}$. Then*

$$M = \sum_{H \subseteq N} w_H^N Z_H Z_H^{\top}$$

Proof. Since $M_{I,J} = w_{I \cup J}$, we have by the change of variables that

$$[M]_{I,J} = \sum_{I \cup J \subseteq H \subseteq N} w_H^N = \sum_{H \subseteq N} \chi_{I \cup J}(H) w_H^N$$

where $\chi_{I \cup J}(H)$ is the 0-1 indicator function such that $\chi_I(H) = 1$ if and only if $I \cup J \subseteq H$. On the other hand, $[Z_H Z_H^{\top}]_{I,J} = [Z_H]_I [Z_H]_J^{\top} = 1$ if $I \cup J \subseteq H$, and 0 otherwise. Therefore $[Z_H Z_H^{\top}]_{I,J} = \chi_{I \cup J}(H)$. \square

By the previous lemma it follows that given a solution by using variables $\{w_I^N\}$ we can obtain a solution with variables $\{w_I : |I| \leq 2t\}$. Viceversa, given any assignment of variables in $\{w_I : |I| \leq 2t\}$ we can find an assignment of variables in $\{w_I^N\}$ such that $M_{I,J} = w_{I \cup J}$ and $M = \sum_{H \subseteq N} w_H^N Z_H Z_H^{\top}$. Indeed, set $w_I^N = 0$ for every I such that $|I| > 2t$. For the remaining ones note that for $|I| \leq 2t$ the square matrix corresponding to the following equalities $w_I = \sum_{I \subseteq H \subseteq N} w_H^N$ is invertible since it is upper triangular.

Lemma 17. [26] *Given $y \in \mathbb{R}^{\mathcal{P}_{2t+2}(N)}$, for the vector $z_I = \sum_{i=1}^n A_{\ell i} y_{I \cup \{i\}} - b_{\ell} y_I$ we have*

$$z_I^N = g_{\ell}(x_I) y_I^N \quad (30)$$

where $g_{\ell}(x_I) = \sum_{i=1}^n A_{\ell i} x_i - b_{\ell}$ is a linear function corresponding to the constraint ℓ , evaluated at x_I such that $x_i = 1$ if $i \in I$ and $x_i = 0$ otherwise.

Proof. We need to show that this choice of z_I^N yields $z_I = \sum_{I \subseteq H \subseteq N} z_H^N$. We plug in (30) to obtain

$$\begin{aligned} \sum_{I \subseteq H \subseteq N} z_H^N &= \sum_{I \subseteq H \subseteq N} g_{\ell}(x_H) y_H^N = \sum_{I \subseteq H \subseteq N} \left[\sum_{i=1}^n A_{\ell i} x_i - b_{\ell} \right]_{x=x_H} y_H^N \\ &= \sum_{I \subseteq H \subseteq N} \left(\sum_{i=1}^n [A_{\ell i} x_i]_{x=x_H} y_H^N - b_{\ell} y_H^N \right) = \sum_{I \subseteq H \subseteq N} \sum_{i=1}^n [A_{\ell i} x_i]_{x=x_H} y_H^N - b_{\ell} y_I \end{aligned}$$

Here the term $[A_{\ell i} x_i]_{x=x_H} y_H^N$ is $A_{\ell i} y_H^N$ if $i \in H$ and 0 otherwise. Taking this into account and changing the order of the sums, the above becomes

$$\sum_{i=1}^n \sum_{I \cup \{i\} \subseteq H \subseteq N} A_{\ell i} y_H^N - b_{\ell} y_I = \sum_{i=1}^n A_{\ell i} y_{I \cup \{i\}} - b_{\ell} y_I$$

which proves the claim. \square

The above discussion together with the observation that $y_\emptyset = 1$ implies that $\sum_{J \subseteq N} y_J^N = 1$ and justifies Definition 1. Finally, we remark the following

Lemma 18. *Let f denote the vector of coefficients of polynomial $f(x)$ of (23). Then the objective value of the solution y is given by*

$$\sum_{I \subseteq N} f_I y_I = \sum_{I \subseteq N} f(x_I) y_I^N$$

Proof. Similar lines as the proof of Lemmas 16 and 17. □

B Change of variables for Max-Cut

Grigoriev [18] and Laurent [27] proved that the following solution is feasible for any $\omega \leq n/2$ up to round $t \leq \lfloor \omega \rfloor$ for the SoS hierarchy given in Definition 4:

$$y_I = \frac{\binom{\omega}{|I|}}{\binom{n}{|I|}} \quad \forall I \subseteq N : |I| \leq 2t$$

Using the change of basis (29), solution $\{y_I\}$ is equivalent to solution $\{y_I^N\}$:

$$\begin{aligned} y_I^N &= \sum_{H \subseteq N \setminus I} (-1)^{|H|} y_{I \cup H} = \sum_{h=0}^{n-|I|} \binom{n-|I|}{h} (-1)^h \frac{\binom{\omega}{|I|+h}}{\binom{n}{|I|+h}} \\ &= y_I \binom{\omega - |I| - 1}{n - |I|} (-1)^{n-|I|} = (n+1) \binom{\omega}{n+1} \frac{(-1)^{n-|I|}}{\omega - |I|} \end{aligned} \tag{31}$$

where we use the identity $\sum_{\omega=0}^m (-1)^\omega \binom{n}{\omega} = (-1)^m \binom{n-1}{m}$.